

**A Bayesian network approach to evaluating
electrical ignition in fire investigation**

by

Daniel Massey

A thesis submitted for the
Degree of Doctor of Philosophy (Science)

University of Technology, Sydney

2005

Certificate of Authorship and Originality

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I received in my research work and the preparation of the thesis itself has been acknowledged. I certify that all information sources and literature are indicated in the thesis.

A handwritten signature in black ink, appearing to read 'Massey', with a stylized flourish extending from the end.

Daniel Massey

August, 2005

Acknowledgements

Credit is due firstly to the supervisors of my postgraduate research, for their continued assistance and friendship over the past few years. My sincere thanks go to Dr Eric du Pasquier, for his strong vision developing the project and sage advice on all things fire-related; to Dr Anne Lear, New South Wales Fire Brigades, for her continued lobbying of fire investigators on my behalf and extremely thorough efforts editing my writing; and to Dr Boris Choy, Department of Mathematical Sciences at UTS, who accepted the role of principal supervisor midway through my project and in doing so saved me from a quagmire of Bayesian statistics and probability. Boris' expertise and motivation was the perfect solution to the unfortunate period I spent without a supervisor and without steady direction.

In a professional capacity, I extend my gratitude to members of the Fire Investigation and Research Unit of the New South Wales Fire Brigade, whose shared knowledge and experience played an important role in the formative stages of my research. Thanks are also due to everyone at the Institut de Police Scientifique at the University of Lausanne, whose hospitality and renowned research pedigree had tremendous influence on the focus of this work and made my time in Switzerland extremely beneficial.

On a personal level, I must thank my fellow forensic postgraduates at UTS, whose parallel suffering and moral support helped maintain perspective even when none of us could see an end in sight. Finally, my heartfelt appreciation to my family and partner for their unwavering daily encouragement, faith and understanding throughout every up and down of the last few years.

Table of Contents

Certificate of Authorship and Originality	ii
Acknowledgements	iii
Table of Contents	iv
List of Figures	ix
List of Tables.....	xiii
Abstract	xiv
1. Fire Science and Fire Investigation.....	1
1.1. Fire science.....	1
1.1.1. Combustion	2
1.1.2. Heat transfer	3
1.1.3. Fuel properties.....	5
1.1.4. Heat release rate	7
1.1.5. Fire growth.....	8
1.2. Fire Investigation	10
1.2.1. The scientific method.....	11
1.2.2. Fire patterns.....	11
1.2.3. Ignition sources	13
1.3. Electrical ignition.....	15
1.4. Electrical ignition statistics	16
1.5. Types of electrical ignition.....	19
1.5.1. Arcing.....	19
1.5.2. Overcurrent due to overload.....	22
1.5.3. Overcurrent due to short circuit or ground fault	23
1.5.4. Poor connections	24
1.5.5. Excess thermal insulation.....	27
1.6. Current Deviation.....	28
1.7. Investigation of electrical ignition	32
1.7.1. Electrical distribution system: Insulation.....	33
1.7.2. Electrical distribution system: Conductors	34
1.7.3. Appliances.....	38

1.7.4.	Instrumental analysis.....	39
1.8.	Aims and Objectives	40
2.	Probability Background	42
2.1.	Basic concepts.....	42
2.1.1.	Terminology	42
2.1.2.	Types of events	42
2.1.3.	Axioms of probability	45
2.1.4.	Further properties of probability	46
2.1.5.	Interpretations of probability.....	47
2.1.6.	Conditional probability	51
2.1.7.	Independence.....	53
2.1.8.	Bayes' Theorem	53
2.1.9.	Odds form of Bayes' Theorem.....	55
2.2.	Bayesian Belief Networks.....	57
2.2.1.	What are Bayesian Belief Networks?	57
2.2.2.	Advantages of Bayesian Networks	65
2.3.	Applications of Bayesian Networks.....	67
3.	Creating a Bayesian network for current deviation.....	71
3.1.	Objectives of the network	71
3.2.	Network design – version 1.....	72
3.2.1.	Hypothesis node	75
3.2.2.	Root node	75
3.2.3.	Branch A	75
3.2.4.	Branch B.....	76
3.2.5.	Branch C.....	77
3.2.6.	Comments on version 1	78
3.3.	Network design – version 2.....	79
3.3.1.	Branch A	80
3.3.2.	Branch B.....	81
3.3.3.	Branch C.....	81
3.3.4.	Comments on version 2.....	82
3.4.	Network design – Final version	82
3.4.1.	Root node	83

3.4.2.	Branch A	84
3.4.3.	Branch B.....	84
3.4.4.	Branch C.....	84
3.4.5.	Comments on final version	85
3.5.	Node definitions	85
3.5.1.	Nodes I, B and E	86
3.5.2.	Nodes IB and CD	88
4.	Types of Bayesian network analysis	90
4.1.	Single-value analysis.....	90
4.1.1.	What is single-value analysis?	90
4.1.2.	Disadvantages of single-value analysis.....	90
4.2.	Distribution-based analysis	91
4.2.1.	Beta distribution	92
4.2.2.	Uniform distribution	99
4.2.3.	Specifying an appropriate beta distribution	100
4.2.4.	Restrictions on input data.....	105
5.	Software Analysis	109
5.1.	Hugin Lite®	109
5.1.1.	Example of single-value analysis with Hugin Lite®	110
5.1.2.	Other uses of Hugin Lite®	111
5.2.	Matlab®	112
5.2.1.	Features and advantages.....	112
5.2.2.	Bayes Net Toolbox.....	114
5.2.3.	Creating the Bayesian network for current deviation with Matlab®	115
5.2.4.	Example of single-value analysis with Matlab®	117
5.2.5.	Distribution-based analysis with Matlab®.....	118
5.3.	Sensitivity analysis.....	127
5.3.1.	Reasons for performing sensitivity analysis	127
5.3.2.	Types of sensitivity analysis	128
5.3.3.	Sensitivity analysis with Matlab®	129
5.3.4.	Example of sensitivity analysis with Matlab®	136
6.	Results and Discussion.....	139
6.1.	Single-value analysis with Hugin Lite®	139

6.2.	Single-value analysis with Matlab®	141
6.3.	Distribution-based analysis with Matlab®.....	145
6.3.1.	Uniform distribution	148
6.3.2.	Effect of the number of repeats.....	150
6.4.	Single sensitivity analysis	152
6.4.1.	Effect of the number of repeats.....	156
6.5.	Double sensitivity analysis.....	160
6.5.1.	Effect of the number of repeats.....	168
6.6.	Statistical analysis of the effect of the number of repeats.....	172
6.7.	Implications for fire investigation.....	176
6.8.	Future work	180
7.	Conclusions	182
8.	Appendices	184
	Appendix 8.1: Equations for α and β	184
	Appendix 8.2: Beta distribution library	185
	8.2.1. Beta distribution with a mean of 0.1	185
	8.2.2. Beta distribution with a mean of 0.2	186
	8.2.3. Beta distribution with a mean of 0.3	187
	8.2.4. Beta distribution with a mean of 0.4	188
	8.2.5. Beta distribution with a mean of 0.5	189
	8.2.6. Beta distribution with a mean of 0.6	190
	8.2.7. Beta distribution with a mean of 0.7	191
	8.2.8. Beta distribution with a mean of 0.8	192
	8.2.9. Beta distribution with a mean of 0.9	193
	Appendix 8.3: Restrictions on input data.....	194
	Appendix 8.4: Single sensitivity analysis with node E.....	195
	Appendix 8.5: Single sensitivity analysis with node I.....	199
	Appendix 8.6: Double sensitivity analysis with nodes I & B	203
	Appendix 8.7: Double sensitivity analysis with nodes E & I	207
	Appendix 8.8: Matlab program: CDEV_FINAL2	211
	Appendix 8.9: Matlab program: BNT_DIST	218
	Appendix 8.10: Matlab program: INPUT_PROMPT2	220
	Appendix 8.11 Matlab program: CDEV_RESULTS.....	230

Appendix 8.12 Matlab program: CDEV_SAVE.....	236
Appendix 8.13 Matlab program: CDEV_LOAD	238
9. Bibliography.....	242

List of Figures

Figure 1.1	Fire tetrahedron	1
Figure 2.1	Venn diagram depicting an event A contained in sample space S	42
Figure 2.2	Event A and its complement, A^c	43
Figure 2.3	Union of events A and B	43
Figure 2.4	Intersection of events A and B	44
Figure 2.5	Mutually exclusive events.....	44
Figure 2.6	The addition rule requires subtraction of the intersection area	46
Figure 2.7	Statistical regularity exhibited by a coin tossing experiment.....	48
Figure 2.9	Reverend Thomas Bayes, 1702-1761	53
Figure 2.10	<i>Left:</i> A directed acyclic graph. <i>Right:</i> A cyclic graph with a feedback cycle from $B \rightarrow C \rightarrow E \rightarrow D \rightarrow B$	59
Figure 2.11	Serial connection.....	62
Figure 2.12	Example of a serial connection	62
Figure 2.13	Diverging connection	63
Figure 2.14	Example of a diverging connection.....	63
Figure 2.15	Converging connection	64
Figure 2.16	Example of a converging connection.....	65
Figure 3.1	Version 1 of the Bayesian network for current deviation	74
Figure 3.2	Extract from version 1 of the network	78
Figure 3.3	Version 2 of the Bayesian network for current deviation	80
Figure 3.4	Final version of the Bayesian network for current deviation.....	83
Figure 3.5	The Bayesian network for current deviation with abbreviated node names.....	86
Figure 4.1	Examples of the beta distribution with their α and β values	96
Figure 4.2	Examples of the beta distribution with their μ and σ values	98
Figure 4.3	When $\alpha = \beta = 1$, the beta distribution is known as the uniform distribution	100
Figure 4.4	The x-axis represents the probability of E_2 and the y-axis is its probability density function	102

Figure 4.5	The x-axis represents the probability of B2 and the y-axis is its probability density function.	103
Figure 4.6	Sample page from the beta distribution reference library. All graphs have a mean of 0.6 and a unique standard deviation.....	104
Figure 4.7	Acceptable combinations of mean and standard deviation.....	107
Figure 4.8	Acceptable mean ranges for the beta distribution, with standard deviation capped at 0.3.....	108
Figure 5.1	The Bayesian network for current deviation in Hugin Lite® 6.4	110
Figure 5.2	The Matlab® command line interface	114
Figure 5.3	Topological order of the network specified in Matlab®.....	116
Figure 5.4	Editing of the M-file is required for the mock case example.....	118
Figure 5.5	Left: Anatomy of a FOR loop. Right: Example of a FOR loop.....	120
Figure 5.6	Pseudocode representation of the FOR loop used in Matlab® Bayesian analysis	120
Figure 5.7	Flow chart representation of the Matlab® program that analyses the Bayesian network for current deviation	122
Figure 5.8	Matlab® GUI for the input of data and selection of the type of analysis.....	124
Figure 5.9	Graphical representations of the prior probability data	126
Figure 5.10	Matlab® GUI with data for a beta distribution analysis of the mock case example	127
Figure 5.11	Combination of two nodes in double sensitivity analysis.....	131
Figure 5.12	Pseudocode representation of the FOR loops used in single sensitivity analysis	132
Figure 5.13	Flow chart depicting the programmatic structure of single sensitivity analysis	133
Figure 5.14	Pseudocode representation of the nested FOR loops used in double sensitivity analysis	134
Figure 5.15	Flow chart depicting the programmatic structure of double sensitivity analysis	135
Figure 5.16	Matlab® GUI with data for a single sensitivity analysis.....	136
Figure 5.17	Flow chart of a single sensitivity analysis for the mock case example	138

Figure 6.1	Hugin Lite® 6.4 after the network has been executed.....	140
Figure 6.2	The updated Bayesian network after node state IB1 has been instantiated	141
Figure 6.3	Matlab® results of the illustrative example prior to entering evidence.....	143
Figure 6.4	Matlab® results of the illustrative example, after instantiation of node state B1	144
Figure 6.5	Matlab® results of a distribution-based Bayesian analysis	146
Figure 6.6	Histogram of X values.....	148
Figure 6.7	Matlab® results of an analysis in which each root node was represented by a uniform distribution	149
Figure 6.8	Results of an analysis with every root node defined by a uniform distribution	150
Figure 6.9	Matlab® screen output of a single sensitivity analysis.....	152
Figure 6.10	Comparison SS1: Sensitivity of the mean of X to the mean of $P(B)$. Error bars depict 95% confidence intervals.....	153
Figure 6.11	Comparison SS2: Sensitivity of the mean of X to the standard deviation of $P(B)$	154
Figure 6.12	Comparison SS3: Sensitivity of the standard deviation of CD2 to the mean of node B	155
Figure 6.13	Comparison SS4: Sensitivity of the standard deviation of X to the standard deviation of node B.....	156
Figure 6.14	Single sensitivity comparisons SS1 (left column) and SS2 (right column) showing the effect of the number of repeats.....	158
Figure 6.15	Single sensitivity comparisons SS3 (left column) and SS4 (right column), showing the effect of the number of repeats.....	159
Figure 6.16	Matlab® screen output from a double sensitivity analysis	161
Figure 6.17	Three dimensional grid with axes labelled x , y and z	162
Figure 6.18	Comparison DS1: Sensitivity of the mean of X to the means of $P(B)$ and $P(E)$	163
Figure 6.19	Modification of graph DS1: wire-frame and 90 degree left rotation	164

Figure 6.20	Comparison DS2: Sensitivity of the mean of X to the standard deviations of P(B) and P(E).	165
Figure 6.21	Comparison DS3: Sensitivity of the standard deviation of X to the means of P(B) and P(E).....	166
Figure 6.22	Comparison DS4: Sensitivity of the standard deviation of X to the standard deviations of P(B) and P(E).....	167
Figure 6.23	Repeat comparisons of DS1 and DS2	170
Figure 6.24	Repeat comparisons of DS3 and DS4	171
Figure 8.1	Sensitivity of the mean of X to the mean of P(E)	195
Figure 8.2	Sensitivity of the mean of X to the standard deviation of P(B)	196
Figure 8.3	Sensitivity of the standard deviation of X to the mean of P(E)	197
Figure 8.4	Sensitivity of the standard deviation of X to the standard deviation of P(E)	198
Figure 8.5	Sensitivity of the mean of X to the mean of P(I)	199
Figure 8.6	Sensitivity of the mean of X to the standard deviation of P(I)	200
Figure 8.7	Sensitivity of the standard deviation of X to the mean of P(I).....	201
Figure 8.8	Sensitivity of the standard deviation of X to the standard deviation of P(I)	202
Figure 8.9	Sensitivity of the mean of X to the means of P(I) and P(B)	203
Figure 8.10	Sensitivity of the mean of X to the standard deviations of P(I) and P(B)	204
Figure 8.11	Sensitivity of the standard deviation of X to the means of P(I) and P(B)	205
Figure 8.12	Sensitivity of the standard deviation of X to the standard deviations of P(I) and P(B)	206
Figure 8.13	Sensitivity of the mean of X to the means of P(I) and P(E).....	207
Figure 8.14	Sensitivity of the mean of X to the standard deviations of P(I) and P(E).....	208
Figure 8.15	Sensitivity of the standard deviation of X to the means of P(I) and P(E).....	209
Figure 8.16	Sensitivity of the standard deviation of X to the standard deviations of P(I) and P(E).....	210

List of Tables

Table 2.1	Example conditional probability table for a root node named A	60
Table 2.2	Conditional probability table for a child node E with two parent nodes, C and D	61
Table 3.1	Conditional probability table for the network extract in Figure 3.2	78
Table 3.2	Conditional probability table for node IB	88
Table 3.3	Conditional probability table for node CD.	89
Table 5.1	Prior probabilities for parent nodes in the illustrative example	111
Table 5.2	Input information provided by an investigator for a Bayesian analysis with the beta distribution.....	126
Table 6.1	Prior probabilities for root nodes in the illustrative example.....	139
Table 6.2	Prior probabilities for root nodes in the illustrative example.....	142
Table 6.3	Prior probability data provided by the investigator	145
Table 6.4	Effect of the number of repeats on various statistical measures in a distribution-based Bayesian analysis	151
Table 6.5	Effect of the number of repeats on the analysis duration of a single sensitivity analysis	157
Table 6.6	Effect of the number of repeats on the analysis duration of a double sensitivity analysis	168
Table 6.7	Hypothesis testing of repeat levels at a margin of 0.05	173
Table 6.8	Hypothesis testing to determine the minimum resolving power of 100 repeats	174
Table 6.9	Hypothesis testing to determine the minimum resolving power of 1,000 repeats	174
Table 6.10	Hypothesis testing to determine the minimum resolving power of 10,000 repeats	175

Abstract

Investigation of the cause and origin of a fire is a task hindered by several sources of uncertainty. Electrical ignition represents a particularly problematic type of fire to investigate, due chiefly to the fragility of the physical evidence and the difficulty distinguishing between electrical damage that has *caused* the fire and electrical damage caused *by* the fire. Current deviation is a type of electrical ignition that is relatively unknown, yet its ability to cause ignition without alerting protective systems makes it more dangerous than a common overcurrent. The lifetime of the process is highly variable, able to proceed without warning anywhere from days to years before ignition. Furthermore, the transient nature of current deviation means that it is especially difficult for investigators to find direct, tangible evidence that it has occurred. Steeped in such a high degree of uncertainty, current deviation was selected as the type of electrical ignition with which to examine a new approach to fire investigation based on Bayesian Belief Networks (BBN).

Bayesian networks are graphical structures that use mathematical probability to represent and analyse influential relationships between variables in a system. This scientifically rigorous method for dealing with uncertainty makes it an ideal aid for investigating current deviation. Further advantages to this approach include the immediate propagation of evidence and the ability of the graphical network to visualise complicated phenomena in an economical and intuitive manner.

A Bayesian network was constructed to represent current deviation, based on the three types of evidence required for ignition. For each type of evidence a fire investigator must express an expert opinion in the form of a numerical probability. Bayesian inference is then used to calculate the likelihood of the hypothesis that conditions existed for current deviation to occur.

Two types of analysis were performed with the Bayesian network for current deviation. A single-value approach using Hugin Lite® software provided a fast and simple method for computing the probability of the hypothesis as a solitary number. The Matlab® software package was then used for an advanced distribution-based

analysis that allowed quantification of uncertainty throughout the investigation. Sensitivity analyses were also implemented to enable the expert to calculate the contribution of each type of evidence and guide the investigation accordingly.

Bayesian networks proved to be an effective decision aid for dealing with uncertainty in the investigation of ignition by current deviation. Recommendations and guidelines for use of this technique were formulated.